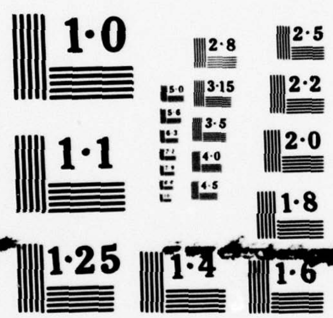
AD-A060 912 STANFORD UNIV CALIF DEPT OF OPERATIONS RESEARCH F/G 15/5
COMPUTATION OF THE OPTIMAL AVERAGE COST POLICY FOR THE TWO TERM--ETC(U)
APR 78 R K DEB N00014-76-C-0418
UNCLASSIFIED TR-78 NL

1 OF 1
ADA
060912

END
DATE
FILMED
1 -79
DDC



NATIONAL BUREAU OF STANDARDS
MICROCOPY RESOLUTION TEST CHART

LEVEL II

12

6 COMPUTATION OF THE OPTIMAL AVERAGE COST POLICY
FOR THE TWO TERMINAL SHUTTLE.

by

10 RAJAT K. DEB

9 Technical rept.

TECHNICAL REPORT NO. 78

11 30 APR[redacted]978

14 TR-78

PREPARED UNDER CONTRACT

15 N00014-76-C-0418, (NR-047-061)

NSF-ENG75-14847

~~FOR THE OFFICE OF NAVAL~~ RESEARCH

Frederick S. Hillier, Project Director

12 35p.

This research was supported in part by National Science Foundation
Grant ENG 75-14847 Department of Operations Research, Stanford University
and issued as Technical Report No. 48

DEPARTMENT OF OPERATIONS RESEARCH

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

COMPUTATION OF THE OPTIMAL AVERAGE COST POLICIES

FOR THE TWO TERMINAL SHUTTLE

Rajat K. Deb*
Stanford University, Stanford, California

In this paper we consider the problem of determining the optimal average cost policy for operating a shuttle between two terminals. The passengers arrive at each of the terminals according to Poisson processes and are transported by a single carrier with capacity $Q \leq \infty$ operating between the terminals. Under a fairly general cost structure, we show that the optimal average cost policy is monotone. Bounds are derived for the optimal control function and computational procedures for determining the optimal policy for both the finite and infinite capacity cases are presented.

* On leave from State University of New York at Oswego. Partially supported by the NSF Grant No. ENG 75-14847 and ONR Grant No. N00014-76-C-0418

COMPUTATION OF THE OPTIMAL AVERAGE COST POLICY
FOR THE TWO TERMINAL SHUTTLE

1. Introduction

In an earlier paper [2] we have shown that a stationary monotone policy minimizes the expected total discounted cost of operating a finite capacity shuttle between two terminals. In this paper we show that the results of [2] can be used to obtain the optimal average cost policies for both the finite capacity and the infinite capacity shuttles. In particular, we present methods by which the optimal policy can actually be computed. For the infinite capacity shuttle the optimal policy can be determined by solving a system of linear equations. However, for the finite capacity case the problem turns out to be much more complex. In Section 4 we present an approximate method for finding the optimal policy for the finite capacity case. The problem of finding these policies is non-trivial because the state space for this problem is infinite. Since this paper is a natural extension of the earlier paper [2], we assume that the reader is familiar with the results of this paper. In the following we briefly describe the model and the various assumptions.

We consider a batch service queue comprising of a carrier with capacity $Q \leq \infty$, operating between two terminals numbered 0 and 1 respectively. Passengers arrive at these terminals according to independent Poisson processes $X(t)$ and $Y(t)$ with respective intensities λ_0 and λ_1. The carrier can be held at these terminals until either a new passenger arrives and another decision is made, or the carrier is dispatched and no decision is made until it arrives at the next terminal. When x passengers are present at the terminal, the batch

size of the passengers boarding the carrier equals $x \wedge Q \equiv \min\{x,Q\}$. The costs associated with operating the system consists of a carrying cost and a holding cost. The cost of carrying y passengers is $R + cy$ and the cost of holding x passengers is hx per unit of time, where R, c and h are nonnegative constants. Without loss of generality, we can assume that no holding cost is charged during the interterminal travel time to the passengers who are already aboard the carrier. The interterminal travel times are assumed to be independent positive random variables with identical distribution $B(\cdot)$, finite mean μ and finite second moment σ^2. Our objective is to determine a policy, that is, a sequence of decision rules which minimizes the long range expected average cost of operating the system. Throughout this paper we assume that $2\lambda_0\mu < Q$ and $2\lambda_1\mu < Q$. Under this assumption, it can be shown that the expected queue length at each terminal is finite. Without loss of generality, we also assume that $c = 0$, because if the expected queue length is finite, then all the arriving customers will be ultimately dispatched and the contribution of the proportional carrying cost to the expected average cost will be the same under all policies. Hence the policy which is optimal for $c = 0$ is also optimal for $c > 0$.

There has been relatively little published work in the area of optimal control of shuttles. In [2] Deb has shown that for the discounted cost case the optimal policy is monotone. He also suggests methods for approximating the optimal discounted policy by linear functions. For the infinite capacity case, Ignall et al. [7] consider the problem of computing the average cost under a simple (not necessarily

optimal) operating policy. In this paper we show that the optimal average cost policy has the following form.

Let the state of the system be denoted by the triplet (x,y,δ) where x and y are the respective numbers of passengers at the terminal 0 and 1 and δ is either 0 or 1 according to whether the carrier is at the terminal 0 or 1 respectively. Then there are monotone decreasing control functions $\tilde{G}_0(y) \leq Q$ and $\tilde{G}_1(x) \leq Q$ such that if $\delta = 0$ $(\delta = 1)$, then the optimal policy is to dispatch the carrier if and only if $x \geq \tilde{G}_0(y)$ $(y \geq G_1(x))$. In Remark 3.8 we show that $\tilde{G}_0(y) \leq m_0 - y$ and $\tilde{G}_1(x) \leq m_1 - x$, where m_0 and m_1 are positive constants. In Section 4.1 we present a policy improvement algorithm and a linear programming formulation for determining the optimal policy for the infinite capacity case. However, for the finite capacity case, it is not easy to compute the optimal policy. If $Q/2\mu$ is considerably larger than the arrival rates λ_0 and λ_1, the results for the infinite capacity case can be used as an approximation to the finite capacity case. However, for small Q the approximation is crude. In Section 4.2 we present a modified policy approximation algorithm for determining the approximate optimal policy for this case.

2. Preliminaries

The notation introduced in this section is used throughout this paper. We let $X(t)$ and $Y(t)$ denote the number of arrivals in time t at the terminal 0 and 1 respectively. Set $Z(t) = X(t) + Y(t)$, $\lambda = \lambda_0 + \lambda_1$ and let the random variables τ, ξ_0 and ξ_1 respectively denote an arbitrary interterminal travel time and arbitrary interarrival times at the terminals 0 and 1. Let $\xi = \min\{\xi_0,\xi_1\}$ and $V_\alpha(x,y,\delta)$

be the optimal α-discounted cost when the system is in state (x,y,δ). Without loss of clarity we shall often suppress the discount factor α. For instance, the α-discounted cost of the policy π is designated by $V_\pi(x,y,\delta)$. We also let ϕ_π and ϕ respectively denote the expected average cost of the policy π and the optimal expected average cost. It can be shown that the optimal α-discounted cost satisfies the following functional equation.

$$(2.1) \qquad V_\alpha(x,y,\delta) = \min\{f(x,y,\delta),\ g(x,y,\delta)\}\ ,$$

where f is the cost of holding the carrier until the next arrival and g is the cost of dispatching the carrier immediately. Letting

$$H(x) = hx/(\alpha+\lambda),\ a = E\{\exp(-\alpha\xi)\} = \lambda/(\alpha+\lambda)\ ,$$

$$p = P[X(\xi) = 1,\ Y(\xi) = 0] = \lambda_0/\lambda\ ,$$

$$q = P[X(\xi) = 0, Y(\xi) = 1] = 1-p = \lambda_1/\lambda,\ \overline{H}(x) = E\int_0^\tau e^{-\alpha t}h(x + Z(t))dt\ ,$$

and

$$d_{ij} = \int_0^\infty e^{-\alpha t}P[X(t) = j,\ Y(t) = i-j]dB(t)\ ,$$

we can write

$$(2.2) \quad f(x,y,\delta) = H(x+y) + apV_\alpha(x+1,y,\delta) + aqV_\alpha(x,y+1,\delta)$$

$$(2.3) \begin{cases} g(x,y,0) = R + \overline{H}(x+y-x \wedge Q) + \sum d_{ij}V_\alpha(x-x \wedge Q+j, y+i-j, 1) \\ g(x,y,1) = R + \overline{H}(x+y-y \wedge Q) + \sum d_{ij}V_\alpha(x+j, y-y \wedge Q+i-j, 1)\ . \end{cases}$$

The summation on d_{ij} is taken over the set $\{i \geq 0,\ 0 \leq j \leq i\}$.

These equations are the same as those derived in [2]. Since we are interested in the average cost in this paper, we can set $\alpha = 0$ in evaluating $H(x)$ and $\overline{H}(x)$ (see p. 161 of [8]). Then for $\alpha = 0$

$$H(x) = hx/\lambda, \ \overline{H}(x) = h\mu x + h\lambda\sigma^2/2 \ . \tag{2.4}$$

In [2] we have shown that (2.1) is well-defined in the sense $V_\alpha < \infty$ and the stationary policy which satisfies (2.1) is optimal. In a fashion similar to equation (2.10) of [2] we define the n-period problem as follows. Let

$$L(x,y,\delta) = R + \overline{H}(x+y-\delta y \wedge Q-(1-\delta)x \wedge Q) \ . \tag{2.5}$$

Set

$$\begin{cases} V^0(x,y,\delta) = (x+y+\lambda/\alpha)h/\alpha \\ V^n(x,y,\delta) = \min\{f^n(x,y,\delta),\ g^n(x,y,\delta)\} \end{cases} \tag{2.6}$$

and for $n \geq 1$

$$f^{n+1}(x,y,\delta) = H(x+y) + apV^n(x+1,y,\delta) + aqV^n(x,y+1,\delta) \tag{2.7}$$

$$g^{n+1}(x,y,\delta) = L(x,y,\delta) + \sum d_{ij}V^n(x-(1-\delta)x \wedge Q+j,y-\delta y \wedge Q+i-j,1-\delta) \ . \tag{2.8}$$

Note that we have suppressed the influence of α on the n-period cost. The function f^n and g^n are the same as defined in (2.2) and (2.3) except that V_α in the right side of (2.2) and (2.3) has been replaced by V^n in (2.7) and (2.8). In [2] we have shown that $V^n \to V$, $f^n \to f$ and $g^n \to g$. In addition, for any function $\omega(x,y,z)$ and $\gamma \in [0,1]$, we define the difference operator Δ as follows:

$$\Delta\omega_\gamma(x,y,z) = \omega(x,y,z) - \omega(x-1+\gamma,y-\gamma,z). \tag{2.9}$$

As before, we sometimes suppress γ. For instance, the statement $\Delta f >$ Δg means $\Delta f_0 > \Delta g_0$ and $\Delta f_1 \geq \Delta g_1$.

In Section 3 we show that there is a discount factor $\tilde{\alpha}$ such that the optimal (stationary) $\tilde{\alpha}$-discounted policy is also optimal for the average cost case. Call this policy π^*. In the remark following Lemma 3.5 we show that π^* exists and is finite. Furthermore, under the policy π^*, the semi-Markov process is positive recurrent. Therefore using arguments similar to those used in the Theorem 7.6 of [8], we can show that ϕ satisfies the following system of equations.

$$v(x,y,\delta) = \min\{\hat{f}(x,y,\delta),\ \hat{g}(x,y,\delta)\}\ , \tag{2.10}$$

where

$$\hat{f}(x,y,\delta) = h\lambda^{-1}(x+y) + pv(x+1,y,\delta) + qv(x,y+1,\delta) - \lambda^{-1}\phi \tag{2.11}$$

and

$$\begin{cases} \hat{g}(x,y,0) = r_0(x,y) - \mu\phi + \sum\limits_{i \geq 0, j \geq 0} p_{ij} v(x-x\wedge Q+i,y+j,1) \\ \hat{g}(x,y,1) = r_1(x,y) - \mu\phi + \sum\limits_{i \geq 0, j \geq 0} p_{ij} v(x+i,y-y\wedge Q+j,0) \end{cases} \tag{2.12}$$

$$\begin{cases} p_{ij} = P[X(\tau) = i,\ Y(\tau) = j] = p_i\tilde{p}_j\ , \\ p_i = P[X(\tau) = i],\ \tilde{p}_j = P[Y(\tau) = j] \\ r_\delta(x,y) = R + h\mu\{x + y - \delta y\wedge Q - (1-\delta)x\wedge Q\} + h\lambda\sigma^2/2\ . \end{cases} \tag{2.13}$$

Note that equations (2.10)-(2.13) can be obtained directly from equations (2.1)-(2.3) by setting $\alpha = 0$ and subtracting $\lambda^{-1}\phi$ and $\mu\phi$

respectively from f and g. The fact that optimal average cost ϕ satisfies (2.10)-(2.13) can be easily checked. Let π be a stationary policy with stationary probabilities $\psi_\pi(x,y,\delta)$ [note that the optimal policy π^* is of the same type]. Let S_π be the set of states such that the carrier is dispatched whenever $(x,y,\delta) \in S_\pi$. Then the average cost

$$\phi_\pi = \{\sum_{S_\pi} \psi_\pi(x,y,\delta) r_\delta(x,y) + \sum_{\tilde{S}_\pi} \psi_\pi(x,y,\delta) h\lambda^{-1}(x+y)\}/\{\lambda^{-1} \sum_{\tilde{S}_\pi} \psi_\pi + \mu \sum_{S_\pi} \psi_\pi\}.$$

Now, suppose ϕ satisfies (2.10); then from (2.10), we have

$$v(x,y,\delta) \leq \begin{cases} \hat{f}(x,y,\delta) & \text{for } (x,y,\delta) \in \tilde{S}_\pi \\ \hat{g}(x,y,\delta) & \text{for } (x,y,\delta) \in S_\pi \end{cases}.$$

Now pre-multiplying both sides of the above inequalities by $\psi_\pi(x,y,\delta)$ and summing over all (x,y,δ), we can show that $\phi \leq \phi_\pi$. In particular, if π^* satisfies (2.10), then the inequalities in the above are replaced by equalities and $\phi = \phi^*$. Also note that the function v inherits the behavior of V_α. In fact, subtracting $V_\alpha(0,0,0)$ from both sides of (2.1) and taking limit as $\alpha \to 0$, we can show that

$$(2.14) \qquad v(x,y,\delta) = \lim_{\alpha \to 0} \{V_\alpha(x,y,\delta) - V_\alpha(0,0,0)\} .$$

This limit exists because the embedded Markov chain for the optimal α-discounted cost is positive recurrent for small α.

3. Average Cost Policy

In this section we extend the results of discounted cost case [2] and show how one can obtain bounds on the control function $\tilde{G}_0$ and

$\tilde{G}_1$. First we show that the optimal average cost is bounded. Consider the policy θ under which the carrier is always dispatched. The resulting queueing system can be analyzed as two separate queues, one at each terminal with mean service time 2μ and respective arrival rates λ_0 and λ_1. Let τ_1 and τ_2 be arbitrary interterminal travel times. Set $T = \tau_1 + \tau_2$; then the random variable T can be viewed as the service time for each of the queues. Note that T has the mean 2μ and second moment $2(\sigma^2 + \mu^2)$. Suppose W_δ is the expected queue length at the terminal δ. Since $2\mu\lambda_\delta < Q$, the queue length W_δ is finite and the average cost $\phi_\theta = R\mu^{-1} + h(W_0 + W_1)$. In particular, if $Q = \infty$, then $W_0 + W_1 = \frac{1}{2}\lambda E\{T^2\}/E\{T\}$ and $\phi_\theta = R\mu^{-1} + h\lambda(\sigma^2+\mu^2)/2\mu$. And if $Q < \infty$, then using (6.1)-(6.8) of [3], the expected queue length is $W_\delta = \lambda_\delta(\sigma^2 + \mu^2)/2\mu + \sum_{i=Q}^{\infty} (i-Q)\pi_i^\delta$, where $\{\pi_i^\delta\}$ are the stationary probabilities of the Markov chain with the transition matrix

$$\beta_{ij} = \begin{cases} \omega_j^\delta & \text{for } i \leq Q \\ \omega_{Q+j-i}^\delta & \text{for } Q < i \leq Q+j \\ 0 & \text{otherwise} \end{cases}$$

where $\omega_j^0 = P[X(T) = j]$ and $\omega_j^1 = P[Y(T) = j]$. Since $2\mu\lambda_\delta < Q$, therefore $W_\delta < \infty$ (See §7 of [9] and §4 of [11]). But ϕ is the optimal average cost and hence

$$\phi < \phi_\theta < \infty \,. \tag{3.1}$$

We summarize the main results of the discounted cost case [2] in Theorems 3.1 and 3.2. These results are then used to show that the optimal average cost policy and the optimal discounted cost policy have the same form as that of the discounted cost case.

Theorem 3.1. If $h > \alpha R/Q$, then

(i) $\Delta V^n \geq \Delta g^n$, $\Delta V^n \leq \Delta f^n$, $\Delta f^n \geq \Delta g^n$

(ii) $\Delta V^n \geq 0$.

This is essentially a restatement of the Theorem 3.3 of [2]. Note that the Theorem 3.1 (i) and (ii) also holds for the infinite time horizon problem (Corollary 4.1 of [2]) and hence

$$(3.2) \qquad \Delta V \geq \Delta g, \quad \Delta V \leq \Delta f, \quad \Delta f \geq \Delta g \quad \text{and} \quad \Delta V \geq 0 .$$

Furthermore, in view of Equation (2.14), the function ν inherits the structure of V_α and therefore

$$(3.3) \qquad \Delta \nu \geq \Delta \hat{g}, \quad \Delta \nu \leq \Delta \hat{f}, \quad \Delta \hat{f} \geq \Delta \hat{g} \quad \text{and} \quad \Delta \nu \geq 0 .$$

Also note that if $\nu(x,y,\delta) = \hat{f}(x,y,\delta)$, then $\hat{f}(x,y,\delta) < \hat{g}(x,y,\delta)$. In addition, either $\nu(x+1,y,0) = \hat{f}(x+1,y,0)$ or $\nu(x+1,y,0) = \hat{g}(x+1,y,0)$. In the first case $\Delta\nu_0(x+1,y,0) = \Delta\hat{f}_0(x+1,y,0) > 0$ and in the second case $\Delta\nu_0(x+1,y,0) > \Delta\hat{g}(x+1,y,0) \geq 0$. Similarly one can show that $\Delta\nu_1(x,y+1,1) > 0$. We use this fact in Theorem 3.7.

Theorem 3.2. If $h > \alpha R/Q$, then there are monotone decreasing functions $G_\delta(\cdot) < Q$, $\delta = 0, 1$, such that following is an optimal α-discounted policy. Suppose the state of the system is (x,y,δ) and $\delta = 0$ $(\delta = 1)$, then the optimal policy is to dispatch the carrier if and only if $x \geq G_0(y)$ $(y \geq G_1(x))$.

This is Theorem 4.2 of [2]. Note that G_δ depends on the discount factor α. In Lemma 3.5 we show that for some $\alpha > 0$, $G_\delta(\cdot)$ is

also an optimal average cost policy. Before doing so we need the following additional lemma and theorem.

For $\alpha > 0$, let $\pi^*(\alpha)$ be the optimal discounted cost policy. Define $\Pi(\beta) = \{\pi^*(\alpha): 0 < \alpha \leq \beta\}$ to be the set of α-optimal policy for each $\alpha \leq \beta$.

Lemma 3.3. If $h > \tilde{\alpha}R/Q$, then $\Pi(\beta)$ is finite for some $\beta > 0$.

Proof: Note that $h > \alpha R/Q$ for all $0 < \alpha \leq \tilde{\alpha}$. Now let θ be the policy under which the carrier is never held at the terminals and let V_θ be the corresponding α-discounted cost. Then $\alpha V_\theta(x,y,\delta) \to \phi_\theta$ as $\alpha \to 0$ and hence for any $\epsilon > 0$, we can choose $\alpha(x,y,\delta)$ such that for all $0 < \alpha \leq \alpha(x,y,\delta)$, $\alpha V_\theta(x,y,\delta) < \phi_\theta + \epsilon$. Now, set $m = 1 +$ integer part of $\{(\phi_\theta + \epsilon)/h\}$, $\tilde{\beta} = \min\{\alpha(x,y,\delta):\ x + y = m,\ \delta = 0, 1\}$ and $\beta = \min(\tilde{\beta},\tilde{\alpha})$. Clearly for all $0 < \alpha \leq \beta$, $x \geq 0$, $y \geq 0$ and $x+y = m$

$$\alpha V_\alpha(x,y,\delta) \leq \alpha V_\theta(x,y,\delta) < \phi_\theta + \epsilon\ . \tag{3.4}$$

Furthermore, $h(x+y) = hm > \phi_\theta + \epsilon$. Now,

$$\begin{aligned} f(x,y,\delta) &= H(x+y) + apV_\alpha(x+1,y,\delta) + aqV_\alpha(x,y+1,\delta) \\ &= V_\alpha(x,y,\delta) + H(x+y) + ap\{V_\alpha(x+1,y,\delta) - V_\alpha(x,y,\delta)\} \\ &\quad + aq\{V_\alpha(x,y+1,\delta) - V_\alpha(x,y,\delta)\} - (1-a)V_\alpha(x,y,\delta)\ . \end{aligned}$$

Using the fact that $\Delta V \geq 0$, $V_\alpha(x,y,\delta) \leq V_\theta(x,y,\delta)$, $\alpha V_\theta \to \phi_\theta$ and $h(x+y) > \phi_\theta + \epsilon$, we have

$$f(x,y,\delta) \geq V_\alpha(x,y,\delta) + H(x+y) - \frac{\alpha}{\alpha+\lambda} V_\theta(x,y,\delta)$$

$$= V_\alpha(x,y,\delta) + \frac{1}{\alpha+\lambda} \{h(x+y) - \alpha V_\theta(x,y,\delta)\}$$

$$> V_\alpha(x,y,\delta) + \frac{1}{\alpha+\lambda} \{h(x+y) - \phi_\theta - \epsilon\} > V_\alpha(x,y,\delta) .$$

Therefore, using (2.1) we have $V_\alpha(x,y,\delta) = g(x,y,\delta)$ and from Theorem 3.2 we obtain $V_\alpha(x',y',\delta) = g(x',y',\delta)$ for all $x' \geq x$ and $y' \geq y$. As a result, for all $\alpha \leq \beta$, $x \geq 0$ and $y \geq 0$, $G_0(y) \leq (\phi_\theta + \epsilon)h^{-1} - y$ and $G_1(x) \leq (\phi_\theta + \epsilon)h^{-1} - x$. Also note that there are only a finite number of these functions G_0 and G_1 because G_δ is a nonnegative integer valued function defined for nonnegative integer values of its argument. Furthermore, a policy is completely specified by the pair (G_0,G_1) and there are only a finite number of such pairs. Therefore, $\Pi(\beta)$ is finite.

Lemma 3.4. For each $\pi \in \Pi(\beta)$, the underlying Markov chain is irreducible.

Proof. In the following we use the notation $x \overset{p}{\rightarrow} y$ to indicate that the probability of transition from state x to state y is at least p. Let π be a policy described in Theorem 3.2 and β be the discount factor defined in Lemma 3.3. Using (2.11), (2.12), Theorem 3.2, Lemma 3.3 and the fact $G_\delta \leq Q$, we have

$$(0,0,0) \xrightarrow{p^{G_0(0)}} (G_0(0),0,0) \xrightarrow{p_{ij}} (i,j,1) \text{ and } (0,Q,1) \xrightarrow{p_{ij}} (i,j,0) .$$

Therefore, $(0,0,0)$ communicates with all other. Now suppose the system is in state (x,y,δ), $x \leq Q$, $Q < y \leq 2Q$ and $\delta = 0$, then

$$(x,y,1) \xrightarrow{p_{00}} (x,y-Q,0) \xrightarrow{p^{G_0(y-Q)-x}} (G_0(y-Q),y-Q,0) \xrightarrow{p_{00}} (0,y-Q,1)$$

$$\xrightarrow{q^{G_1(0)-y+Q}} (0,G_1(0),1) \xrightarrow{p_{00}} (0,0,0) .$$

In a similar fashion, we can show that for all $x \geq 0$ and $y \geq 0$, (x,y,δ) communicates with $(0,0,0)$. Therefore, the embedded Markov chain is irreducible.

Under each $\pi \in \Pi(\beta)$, the resulting queueing process can be analyzed by the embedded Markov chain or semi-Markov process. Since the embedded Markov chain is irreducible, the semi-Markov process is either positive recurrent or null recurrent (transient). In the first case, by the strong law of large numbers for these processes, the average queue length is finite and the average cost $\phi_\pi < \infty$. In the second case the expected queue length is infinite and so is the average cost ϕ_π. In either case by abelian theorem $\alpha V_\pi \to \phi_\pi$ as $\alpha \to 0$, where $+\infty$ can be included as a possible limit (Lemma 5.1 of [3]).

Furthermore, since $\Pi(\beta)$ is finite, there is a $\pi^* \in \Pi(\beta)$ and a sequence of discount factors $\{\alpha_n\}$, $\alpha_n \to 0$, such that π^* is α_n-optimal.

Lemma 3.5. If $h > 0$, then π^* is an optimal average cost policy.

Proof. Since $h > 0$, we can find a $\tilde{\alpha}$ such that $h > \tilde{\alpha}R/Q$ and the hypothesis of Lemma 3.3 is satisfied. Now, for any policy π (not necessarily in $\Pi(\beta)$), using Theorem 1 on p. 181 of [12], we get

$$\phi_\pi \geq \overline{\lim_{\alpha \to 0}}\ \alpha V_\pi \geq \overline{\lim_{n \to \infty}}\ \alpha_n V_{\pi(\alpha_n)} = \overline{\lim_{n \to \infty}}\ \alpha_n V_{\pi^*}$$

$$= \lim_{\alpha \to 0} \alpha V_{\pi^*} = \phi^* .$$

This completes the proof.

Remark: From (3.1) we know $\phi_{\pi^*} < \infty$. Since the embedded Markov chain is irreducible, therefore the resulting semi-Markov process is positive recurrent.

Theorem 3.6. If $h > 0$, then there are monotone decreasing functions $\tilde{G}_\delta < Q$, $\delta = 0, 1$, such that if $\delta = 0$ $(\delta = 1)$ and the state of the system is (x,y,δ), then it is optimal to dispatch if and only if $x \geq \tilde{G}_0(y)$ $(y \geq \tilde{G}_1(x))$.

Proof. Clearly π^* as defined in Lemma 3.5 is optimal. But π^* is also an α_n-optimal policy and hence using Theorem 3.2 we obtain the desired result.

The following theorem sharpens the bounds on G_0 and G_1 developed in the Lemma 3.3.

Theorem 3.7. For any $\delta \in \{0,1\}$, let

$$(3.5) \quad \begin{cases} S_0(y) = h^{-1}\phi - \{y + \lambda_1\mu + h^{-1}\lambda_1 \sum p_{ij}\Delta v_1(i,y+1+j,1)\} \\ S_1(x) = h^{-1}\phi - \{x + \lambda_0\mu + h^{-1}\lambda_0 \sum p_{ij}\Delta v_0(x+1+i,j,1)\} . \end{cases}$$

Then

$$(3.6) \qquad \tilde{G}_\delta(\cdot) = \min\{S_\delta(\cdot),Q\} .$$

Proof. Suppose $\delta = 0$ and $x < \tilde{G}_0(y)$, then it suffices to show that $v(x,y,0) = \hat{f}(x,y,0)$. Suppose the assertion is false and $v(x,y,0) = \hat{g}(x,y,0)$. Then from (3.6) and Theorem 3.6, we have $x \leq Q-1$ and $v(x',y',0) = \hat{g}(x',y',0)$ for all $x' \geq x$, $y' \geq y$. Now using (2.10)-(2.12), we obtain

$$\begin{aligned}
\hat{f}(x,y,0) &= h\lambda^{-1}(x+y) + pv(x+1,y,0) + qv(x,y+1,0) - \phi\lambda^{-1} \\
&= v(x,y,0) + h\lambda^{-1}(x+y) + p\{v(x+1,y,0) - v(x,y,0)\} \\
&\quad + q\{v(x,y+1,0) - v(x,y,0)\} - \phi\lambda^{-1} \\
&= v(x,y,0) + h\lambda^{-1}(x+y) + q\{h\lambda\mu + \textstyle\sum p_{ij}v(i,y+1+j,1)\} - \lambda^{-1}\phi \\
&= v(x,y,0) + h\lambda^{-1}\{x-[h^{-1}\phi-y-\lambda_1\mu-h^{-1}\lambda_1 \textstyle\sum p_{ij}v(i,y+1+j,1)]\} \\
&= v(x,y,0) + h\lambda^{-1}[x - S_0(y)] < v(x,y,0) .
\end{aligned}$$

But this contradicts (2.10) and hence $v(x,y,0) = \hat{f}(x,y,0)$. Now to prove the converse, suppose that $v(x,y,0) = \hat{f}(x,y,0)$. Then clearly $x \leq Q-1$ and

$$\begin{aligned}
\hat{f}(x,y,0) &= h\lambda^{-1}(x+y) + pv(x+1,y,0) + qv(x,y+1,0) \\
&= v(x,y,0) + h\lambda^{-1}(x+y) + p\{v(x+1,y,0) - v(x,y,0)\} \\
&\quad + q\{v(x,y+1,0) - v(x,y,0)\} - \lambda^{-1}\phi .
\end{aligned}$$

But from the remark following Theorem 3.1, we know that $\Delta v_0(x,y+1,0) \geq \Delta\hat{g}_0(x,y+1,0)$ and $\Delta v_0(x+1,y,0) > 0$. Therefore

$$\begin{aligned}
\hat{f}(x,y,0) - v(x,y,0) &> h\lambda^{-1}(x+y) + q\{h\lambda\mu + \textstyle\sum p_{ij}v(i,y+1+j,1)\} - \lambda^{-1}\phi \\
&= h\lambda^{-1}[x - S_0(y)] .
\end{aligned}$$

But $\hat{f}(x,y,0) - v(x,y,0) = 0$ and hence $x < S_0(y)$. This completes the proof for $\delta = 0$. For $\delta = 1$ the proof is essentially the same as in the above.

Also note that the above theorem could be obtained directly from Equations (2.1)-(2.3). In this case the proof is similar to that of Lemma 3.3.

Remark 3.8: For any policy π, let $m_\delta(i) = h^{-1}\phi_\pi - (i+\lambda_{1-\delta}u)$, then $m_\delta \geq \tilde{G}_\delta$. This follows from the fact that $\Delta v \geq 0$ and hence using (3.5) and (3.6), we obtain

$$m_\delta - \tilde{G}_\delta = h^{-1}(\phi_\pi - \phi) + \sum p_{ij}\Delta v(\cdot,\cdot,1-\delta) \geq 0 .$$

The following lemma is valid for the special case $Q = \infty$.

Lemma 3.9. Let $Q = \infty$, then

(i) $\Delta^2 V_0(x,y,0) \equiv \Delta V_0(x,y,0) - \Delta V_0(x-1,y,0) \leq 0$

(ii) $\Delta^2 V_1(x,y,1) \leq 0$

Proof. The proof is by induction using the finite period problem defined in (2.6). First we show that if $\Delta^2 f_\delta^k(x,y,\delta) \leq 0$ and $\Delta^2 g_\delta^k(x,y,\delta) = 0$, then $\Delta^2 V_\delta^k(x,y,\delta) \leq 0$. Suppose $\delta = 0$, then from (2.6) we know that either $V^k(x-1,y,0) = f^k(x-1,y,0)$, or $V^k(x-1,y,0) = g^k(x-1,y,0)$. In the first case

$$\Delta V_0^k(x,y,0) \leq \Delta f_0^k(x,y,0) \quad \text{and} \quad \Delta V_0^k(x-1,y,0) \geq \Delta f_0^k(x-1,y,0)$$

and hence

$$\Delta^2 V_0^k(x,y,0) \leq \Delta f_0^k(x,y,0) - \Delta f_0^k(x-1,y,0) = \Delta^2 f_0^k(x,y,0) \leq 0 .$$

In the second case

$$\Delta V_0^k(x,y,0) \leq \Delta g_0^k(x,y,0), \ \Delta V_0^k(x-1,y,0) \geq \Delta g_0^k(x-1,y,0)$$

and hence

$$(3.7) \qquad \Delta^2 V_0^k(x,y,0) \leq \Delta^2 g_0^k(x,y,0) = 0 .$$

Similarly we can show that

$$(3.8) \qquad \Delta^2 V_1^k(x,y,1) \leq 0 .$$

Furthermore from (2.6)-(2.8) we have for all $k \geq 1$, $\Delta^2 g_\delta^k(x,y,0) = 0$ and for $k = 1$,

$$(3.9) \qquad \Delta^2 f_\delta^k(x,y,\delta) = \Delta f_\delta^1(x,y,\delta) - \Delta f_\delta^1(x,y,\delta) = 0 .$$

We now show that $\Delta^2 f_\delta^k(x,y,\delta) \leq 0$ for all $k \geq 1$. Assume (3.9) to be true for all $k \leq n$. Then (3.7) and (3.8) are true for all $k \leq n$. Then from the definition of $f(x,y,\delta)$, we have

$$\Delta^2 f^{n+1}(x,y,\delta) = ap\Delta^2 V_\delta^n(x+1,y,\delta) + aq\Delta^2 V_\delta^n(x,y+1,\delta) \leq 0 .$$

Therefore, $\Delta^2 V_\delta^n(x,y,\delta) \leq 0$ for all $n \geq 1$. Since $V^n \to V$ as $n \to \infty$, we get $\Delta^2 V_\delta^n \to \Delta^2 V_\delta$ and $\Delta^2 V_\delta(x,y,\delta) \leq 0$.

<u>Remark 3.10</u>: As a consequence of Lemma 3.9, we immediately conclude $\Delta^2 v_\delta(x,y,\delta) \leq 0$. Now using (3.5) and (3.6) one can show that $S_0(y-1) - S_0(y) \leq 1$ and $S_1(x-1) - S_1(x) \leq 1$. Therefore if $\nu(x,y,0) = \hat{g}(x,y,0)$ $(\nu(x,y,1) = \hat{g}(x,y,1))$, then $\nu(x+1,y-1,0) = \hat{g}(x+1,y-1,0)$ $(\nu(x-1,y+1,1) = \hat{g}(x-1,y+1,1))$.

4. Computation of Optimal Average Cost Policy

In this section we develop algorithms for computing the control functions $\tilde{G}_0$ and $\tilde{G}_1$. We treat the cases $Q = \infty$ and $Q < \infty$ separately. Our main tool for developing these algorithm is the system of equations (2.10)-(2.13).

1. Case $Q = \infty$.

Let θ be the policy described in Lemma 3.3; then using the discussion at the beginning of Section 3, the average cost of the policy $\phi_\theta = \{2R + h\lambda(\sigma^2 + \mu^2)\}/2\mu$. From Remark 3.8, we know

$$\tilde{G}_0(y) \leq h^{-1}\phi_\theta - \lambda_1\mu - y \quad \text{and} \quad \tilde{G}_1(x) \leq h^{-1}\phi_\theta - \lambda_0\mu - x \, . \tag{4.1}$$

Set $M = [h^{-1}\phi_\theta - \lambda_1\mu]$ and $N = [h^{-1}\phi_\theta - \lambda_0\mu]$, where the closed brackets denote the integer larger than or equal to the number within the brackets. Then from (4.1), it follows that for $\delta = 0$ $(\delta = 1)$ and $x+y \geq M$ $(x+y \geq N)$, $\nu(x,y,\delta) = \hat{g}(x,y,\delta)$. Furthermore, we assume that $M \geq N$, otherwise we renumber the terminals accordingly. Using (2.10)-(2.13) and writing $\tilde{\delta} = (1-\delta)$ and $r = R + h\lambda\sigma^2/2$, the optimal average ϕ satisfies the following functional equation

$$\begin{aligned} \nu(x,y,\delta) = \min\{&h\lambda^{-1}(x+y) + p\nu(x+1,y,\delta) + q\nu(x,y+1,\delta) - \lambda^{-1}\phi \, , \\ &r - \mu\phi + h\mu[x\delta+\tilde{\delta}y] + \sum_{i \geq 0, j \geq 0} p_i\tilde{p}_j\nu(i+x\delta, j+y\tilde{\delta}, \tilde{\delta})\} \, . \end{aligned} \tag{4.2}$$

Now, using definition M, N and $\hat{g}$ and Theorem 3.7, we obtain the following equalities

$$(4.3)\quad \begin{cases} \nu(x,y,0) = \nu(M-y,y,0) & \text{for } x+y \geq M \geq y+1 \\ \qquad\qquad = \nu(0,y,0) & \text{for } y \geq M \\ \nu(x,y,1) = \nu(x,N-x,1) & \text{for } x+y > N \geq x+1 \\ \qquad\qquad = \nu(x,0,1) & \text{for } x \geq N \end{cases}$$

$$(4.4)\quad \begin{cases} \nu(x,y,0) = (y-N)h\mu + \nu(x,N,0) & \text{for } y \geq N \\ \qquad\qquad = (y-N)h\mu + \nu(M-N,N,0) & \text{for } y \geq N,\ x \geq M-N \\ \nu(x,y,1) = (x-M)h\mu + \nu(M,0,1) & \text{for } x \geq M\ . \end{cases}$$

Using (4.3) and (4.4), we reduce the infinite state minimization problem (4.2) into a finite state problem. Let

$$(4.5)\quad R(x) = \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} p_i \tilde{p}_j \nu(x+i,j,0) \quad \text{and} \quad \tilde{R}(y) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} p_i \tilde{p}_j \nu(i,y+j,1)\ .$$

Then by repeated use of (4.3) and (4.4) and lengthy algebraic manipulation, we have for $x \leq M-1$

$$(4.6)\quad R(x) = \sum_{j=0}^{M-x-1} \tilde{p}_j \sum_{i=0}^{M-x-j-1} p_i \nu(x+i,j,0) + \sum_{j=0}^{M-x-1} \tilde{p}_j P(M-x-j)\nu(M-j,j,0)$$

$$+ \sum_{j=M-x}^{M} \tilde{p}_j \nu(M-j,j,0) + \sum_{j=M+1}^{\infty} \tilde{p}_j (j-M)h\mu + \tilde{P}(M+1)\nu(0,M,0)$$

where

$$(4.7)\quad P(i) = \sum_{k=i}^{\infty} p_k \quad \text{and} \quad \tilde{P}(i) = \sum_{k=i}^{\infty} \tilde{p}_k\ .$$

And for $x \geq M$

$$(4.8)\quad R(x) = \sum_{j=0}^{M} \tilde{p}_j \nu(M-j,j,0) + \sum_{j=M+1}^{\infty} \tilde{p}_j (j-M)h\mu + \tilde{P}(M+1)\nu(0,M,0)\ .$$

Similarly for $y \leq N-1$, we obtain

$$(4.9) \quad \tilde{R}(y) = \sum_{i=0}^{N-y-1} p_i \sum_{j=0}^{N-y-i-1} \tilde{p}_j \nu(i,y+j,0) + \sum_{i=0}^{N-y-1} p_i \tilde{P}(N-y-i)\nu(i,N-i,1)$$

$$+ \sum_{i=N-y}^{N} p_i \nu(i,N-i,1) + \sum_{i=N+1}^{M} p_i \nu(i,0,1)$$

$$+ \sum_{i=M+1}^{\infty} p_i (i-M)h\mu + P(M+1)\nu(M,0,1) ,$$

and for $y \geq N$

$$(4.10) \quad \tilde{R}(y) = \sum_{i=0}^{N} p_i \nu(i,N-i,1) + \sum_{i=N+1}^{M} p_i \nu(i,0,1) + \sum_{i=M+1}^{\infty} p_i (i-M)h\mu$$

$$+ P(M+1)\nu(M,0,1) .$$

Note that the infinite sums $R(x)$ and $\tilde{R}(y)$ depend on $\nu(i,j,0)$, $i+j \leq M$; $\nu(i,j,1)$, $i+j \leq N$ and $\nu(i,0,1)$ for $i = N+1, N+2, \ldots, M$. Furthermore, the value of $R(x)$ and $\tilde{R}(y)$ are respectively independent of x and y for $x \geq M$ and $y \geq N$. We can also express $\tilde{R}(y)$ in terms of $R(x)$ as follows. First note that

$$(4.11) \quad \tilde{R}(y) = r + h\mu^2\lambda_0 - \mu\phi + \sum_{i=0}^{M-1} p_i R(i) + P(M)R(M) \quad \text{for } y \geq N.$$

Then using (4.9), we have

$$(4.12) \quad \tilde{R}(y) = \sum_{i=0}^{N-y-1} p_i \sum_{j=0}^{N-y-i-1} \tilde{p}_j \nu(i,y+j,0) + \sum_{i=0}^{N-y-1} p_i \tilde{P}(N-y-i)\nu(i,N-i,1)$$

$$+ \sum_{i=N-y}^{\infty} p_i [r + \mu hi - \mu\phi + R(i)] .$$

In (4.11) and (4.12) $\tilde{R}(y)$ depend on $\nu(i,j,1)$, $i+j \le N$, a reduction of $(M-N)$ variables from (4.9) and (4.10). The minimization problem (4.2) is then clearly equivalent to the following linear programming problem

Maximize ϕ
subject to

$$\nu(x,y,\delta) - p\nu(x+1,y,\delta) - q\nu(x,y+1,\delta) + \lambda^{-1}\phi \le h\lambda^{-1}(x+y) \tag{4.13}$$

$$\text{for } \delta \in \{0,1\},\ x+y \le M\tilde{\delta} + N\delta - 1$$

$$\nu(x,y,\delta) - \sum p_i\tilde{p}_j\nu(i+x\delta,j+y\tilde{\delta}) + \mu\phi \le r + h\mu(x\delta+y\tilde{\delta}) \tag{4.14}$$

$$\text{for } \delta \in \{0,1\},\ x+y \le M\tilde{\delta} + N\delta - 1$$

$$\nu(x,y,\delta) - \sum p_i\tilde{p}_j\nu(i+x\delta,j+y\tilde{\delta}) + \mu\phi = r + h\mu(x\delta+y\tilde{\delta}) \tag{4.15}$$

$$\text{for } \delta \in (0,1),\ x+y = M\tilde{\delta} + N\delta$$

$$\nu(x,y,\delta) \ge 0 .$$

If we replace $\sum p_i\tilde{p}_j\nu(i+x\delta,j+y\tilde{\delta})$ in (4.14) and (4.15) by right-hand side of (4.6)-(4.10), the above linear program has at most $M(M+1) + N(N+1) + (N-N)$ variables. The fact that this linear program (4.13)-(4.15) indeed gives the optimal average cost follows from remarks at the end of Equations (2.10)-(2.13). The optimal policy $\tilde{G}_0(y)$ $(\tilde{G}_1(x))$ is given by the smallest value of $x(y)$ for given values of $y(x)$, for which equality holds in the above inequality (4.14) and (4.15). Since M and N depend on ϕ_θ, the number of variables in (4.6)-(4.10) can

be reduced by choosing a policy π such that $\phi_\pi < \phi_\theta$ and using ϕ_π in (4.1) for computing M and N. However, it may be difficult to find a policy π such that $\phi_\pi < \phi_\theta$. Moreover, the computation of ϕ_π for a given policy π is itself a non-trivial task. In the following we describe a policy improvement algorithm, which takes advantage of the structure of $\nu(x,y,\delta)$ and the value of ϕ at each iteration of the algorithm. Let $\{\pi(i)\}$ be a sequence of policies and let $S_{\pi(i)}$ be the set of states under the policy $\pi(i)$ such that the carrier is dispatched whenever $(x,y,\delta) \in S_{\pi(i)}$. In addition, we assume that for all $x \geq 0$ and $y \geq 0$, such that $x+y \geq M\tilde{\delta}+N\delta$, $(x,y,\delta) \in S_{\pi(i)}$. Clearly the sequence $\{\pi(i)\}$ contains the optimal policy π^*. Then the policy improvement algorithm consists of the following steps.

1. Set $i = 0$, $S_{\pi(i)} = S_\pi = \{(x,y): x \geq 0, y \geq 0\}$, compute $R(x)$ and $\tilde{R}(y)$
2. Solve (4.16) for ϕ_π and $\nu(x,y,\delta)$

$$(4.16)\quad \begin{cases} \nu(x,y,\delta) + \lambda^{-1}\phi_\pi = h\lambda^{-1}(x+y) + p\nu(x+1,y,\delta) + q\nu(x,y+1,\delta) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{for } (x,y,\delta) \in \tilde{S}_{\pi(i)} \\ \nu(x,y,0) + \mu\phi_\pi = r + h\mu y + \tilde{R}(y) \quad \text{for } (x,y,0) \in S_{\pi(i)} \\ \nu(x,y,1) + \mu\phi_\pi = r + h\mu x + R(x) \quad \text{for } (x,y,0) \in S_{\pi(i)} \end{cases}$$

where $\tilde{S}_\pi$ is the compliment of the set S_π.

Note that the system of equations (4.16) is over-determined and hence we can choose one of the variable ν arbitrarily (set $\nu(0,0,0) = 0$). The total number of unknown variables is $\frac{1}{2}\{M(M+1) + N(N+1)\}$ which is half of the number of variables in the linear programming formulation.

3. Compute S_π , M, and N using the following relationship.

(4.17) $M = [h^{-1}\phi_\pi - \lambda_1\mu]$, $N = [h^{-1}\phi_\pi - \lambda_0\mu]$ and $S_\pi = \{(x,y,\ ): x+y \geq M\tilde{\delta} + N\delta\}$.

4. For all $(x,y,\delta) \in \tilde{S}_\pi$, compute the test quantities t_0 and t_1 using

$$(4.18) \qquad t_0 = \begin{cases} \mu^{-1}[r + h\mu y + \tilde{R}(y) - \nu(x,y,\delta)] & \text{for } \delta = 0 \\ \mu^{-1}[r + h\mu x + R(x) - \nu(x,y,\delta)] & \text{for } \delta = 1 \end{cases}$$

$$(4.19) \qquad t_1 = \lambda[h\lambda^{-1}(x+y) + p\nu(x+1,y,\delta) + q\nu(x,y+1,\delta) - \nu(x,y,\delta)] .$$

If $t_0 < t_1$ for some (x,y,δ), then set $S_\pi = S_\pi \cup (x,y,\delta)$.

5. If $S_{\pi(i)} = S_\pi$ then stop and the optimal action is determined by S_π. Otherwise, set $i=i+1$, $S_{\pi(i)} = S_\pi$, recompute $R(x)$ and $\tilde{R}(y)$ for the new values of M and N. Then go to step 2 of this algorithm.

Note that at each iteration the value of M and N decreases and hence the number of unknown variables in (4.16) reduces. Furthermore, in Step 4 of the algorithm, we have to compute only one of the test quantities t_0 and t_1, because if $(x,y,\delta) \in S_{\pi(i)}$, then $t_0 = \phi_\pi$, otherwise $t_1 = \phi_\pi$.

Also considerable computational savings can be achieved by taking advantage of the structure of ν. For instance, if $\Delta\nu \geq \Delta\hat{g}$, then one can show that if $t_0 < t_1$ for some (x,y,δ), then for all $x' \geq x$ and $y' \geq y$, $S_\pi = S_\pi \cup (x',y',\delta)$. It seems that for the policy improvement algorithm the assertion that $\Delta\nu \geq \Delta\hat{g}$ at each iteration is true. However, we are unable to prove this.

2. Case $Q < \infty$.

The computation of the optimal policy for the case $Q < \infty$ is more complex than the infinite capacity case. When $Q < \infty$, it is not possible to obtain equalities of the form (4.3) and (4.4) and hence the system of Equations (2.10)-(2.13) cannot be expressed as a system finite number of equations. However, one could use the solution of the case $Q = \infty$ as an approximation. Another alternative would be to allow finite waiting space at each terminal. In the following we suggest yet another approximate method.

Let θ be the policy described at the beginning of Section 3, then the average cost of the policy θ is

$$\phi_\theta = R\mu^{-1} + h\lambda(\sigma^2 + \mu^2)/2\mu + h\sum_{i=Q}^{\infty} (Q-i)(\pi_i^0 + \pi_i^1) \;. \tag{4.20}$$

Set

$$M = [h^{-1}\phi_\theta - \lambda_1\mu] \quad \text{and} \quad N = [h^{-1}\phi_\theta - \lambda_0\mu] \;. \tag{4.21}$$

From the Remark 3.8, it follows that if $x+y \geq M\tilde{\delta} + N\delta$, then for all $x' \geq x$ and $y' \geq y$, $v(x,y,\delta) = \hat{g}(x,y,\delta)$. Now using Theorem 3.7, Remark 3.8, (2.10)-(2.13) and writing $r = R + h\lambda\sigma^2/2$, the optimal average ϕ satisfies the following system of equations.

$$\begin{aligned} v(x,y,\delta) = \min\{&h\lambda^{-1}(x+y)+pv(x+1,y,\delta)+qv(x,y+1,\delta)-\lambda^{-1}\phi\;, \\ &r-\mu\phi+h\mu(x+y-\delta y\wedge Q-\tilde{\delta}x\wedge Q) + \sum p_{ij}v(x-\delta x\wedge Q+i,y-\tilde{\delta}y\wedge Q+j,\tilde{\delta})\} \end{aligned}$$

$$\text{for all} \begin{cases} x+y \leq M-1,\ x \leq Q-1 & \text{if } \delta = 0 \\ x+y \leq N-1,\ y \leq Q-1 & \text{if } \delta = 1 \end{cases} \tag{4.22}$$

$$v(x,y,\delta) = r-\mu\phi+h\mu(x+y-\delta y \quad Q-\tilde{\delta}x \quad Q) + \sum p_{ij}v(x-\delta x \quad Q+i,y-\tilde{\delta}y \quad Q+j,\tilde{\delta})$$

otherwise .

Note that (4.22) has infinitely many variables. We shall now reduce the number of variables by approximating $v(x,y,\delta)$ for large values of x and y. Suppose $\delta = 0$, $x \gg Q$ and $y \gg Q$; then using Theorem 3.7, we know that the carrier will not wait at any terminal until the number of waiting passengers at each terminal is less than Q. Let the random variables T_0 and T_1 be the instants at which the number of passengers at terminals 0 and 1 respectively fall below Q for the first time given that initially the system is in state $(x,y,0)$. Then the policy θ is optimal for the duration $\min(T_0,T_1)$ and under this policy

$$(4.23) \qquad \begin{cases} v(x,y,0) - v(x-1,y,0) \approx hE\{T_0\} \\ v(x,y,0) - v(x,y-1,0) \approx hE\{T_1\} . \end{cases}$$

Now, suppose n_0 and n_1 are the respective number of dispatches in time T_0 and T_1. Since Q passengers are removed from each terminal at each dispatch, hence

$$(4.24) \qquad \begin{cases} -QE\{n_0\} + \lambda_0 E\{T_0\} + x \simeq Q \\ -QE\{n_1\} + \lambda_1 E\{T_1\} + y \simeq Q . \end{cases}$$

Also note that

$$(4.25) \qquad \begin{cases} E\{T_0\} = 2\mu E\{n_0\} \\ E\{T_1\} = \mu + 2\mu E\{n_1\} . \end{cases}$$

Then from (4.24) and (4.25) we have

$$(4.26)\qquad \begin{cases} E\{n_0\} \simeq (x-Q)/(Q-2\mu\lambda_0) \\ E\{n_1\} \simeq (y+\mu\lambda_1 - Q)/(Q-2\mu\lambda_1)\ . \end{cases}$$

Using (4.23), (4.26) and (4.25), we obtain

$$(4.27)\qquad \begin{cases} \nu(x,y,0) - \nu(x-1,y,0) \simeq h\rho_0(x-Q) \\ \nu(x,y,0) - \nu(x,y-1,0) \simeq h\rho_1(y-Q/2) \end{cases}$$

where

$$(4.28)\qquad \rho_\delta = 2\mu/(Q-2\mu\lambda_\delta) \quad \text{for} \quad \delta = 0, 1\ .$$

In a similar fashion, we can show that

$$(4.29)\qquad \begin{cases} \Delta\nu_0(x,y,1) \simeq h\rho_0(x-Q/2) \\ \Delta\nu_1(x,y,1) \simeq h\rho_1(y-Q)\ . \end{cases}$$

Suppose $D \gg Q$, then using the approximation (4.27), (4.28), (4.7) and (2.13), we have

$$\begin{aligned} R_0(x,y) &\equiv \sum p_i \sum \tilde{p}_j \nu(x+i,y+j,0) \\ &= \sum_{i=0}^{D-x-1} p_i \sum_{j=0}^{D-y-1} \tilde{p}_j \nu(x+i,y+j,0) \\ &\quad + h\rho_0 \tilde{P}(D-y) \sum_{i=0}^{\infty} p_{D-x+i}\, i\{D-Q+(i+1)/2\} \\ &\quad + h\rho_1 P(D-x) \sum_{j=0}^{\infty} \tilde{p}_{D-y+j}\, j\{D+(j+1-Q)/2\} \\ &\quad + \tilde{P}(D-y)P(D-x)\nu(D,D,0) \end{aligned} \qquad (4.30)$$

and

$$R_1(x,y) \equiv \sum p_i \sum \tilde{p}_j v(x+i,y+j,1)$$

$$= \sum_{i=0}^{D-x-1} p_i \sum_{j=0}^{D-y-1} \tilde{p}_j v(x+i,y+j,1)$$

$$+ h\rho_0 \tilde{P}(D-y) \sum_{i=0}^{\infty} p_{D-x+i}\, i\{D+(i+1-Q)/2\} \tag{4.31}$$

$$+ h\rho_1 P(D-x) \sum_{j=0}^{\infty} \tilde{p}_{D-y+j}\, j\{D-Q+(i+1)/2\}$$

$$+ \tilde{P}(D-y)P(D-x)v(D,D,1) \; .$$

Note that both R_0 and R_1 depend on $v(x,y,\delta)$, $x \leq D$ and $y \leq D$. As before, the first term in R_δ vanishes for $x = D$ and $y = D$.

Now, using (4.30) and (4.31), we can rewrite the system of equations (4.22) as follows. Set

$$\begin{cases} \tilde{S}_\pi = \{(x,y,0)\colon x+y \leq M-1,\ x \leq Q-1\} \cup \{(x,y,1)\colon x+y \leq N-1,\ y \leq Q-1\} \\ S_\pi = \{(x,y,\delta)\colon x \leq D,\ y \leq D \text{ and } (x,y,\delta) \notin \tilde{S}_\pi\} \; . \end{cases} \tag{4.32}$$

Then

$$\begin{cases} v(x,y,\delta) = \min\{h\lambda^{-1}(x+y) - \lambda^{-1}\phi + pv(x+1,y,\delta) + qv(x,y+1,\delta) \; , \\ \qquad r - \mu\phi + h\mu(x+y-\delta y \wedge Q-\tilde{\delta}x \wedge Q) + R_\delta(x-\delta x \wedge Q, y-\tilde{\delta}y \wedge Q)\} \\ \qquad\qquad \text{for } (x,y,\delta) \in \tilde{S}_\pi \\ v(x,y,\delta) = r - \mu\phi + h\mu(x+y-\delta y \wedge Q-\tilde{\delta}x \wedge Q) + R_\delta(x-\delta x \wedge Q, y-\tilde{\delta}y \wedge Q) \\ \qquad\qquad \text{for } (x,y,\delta) \in S_\pi \; . \end{cases} \tag{4.33}$$

Note that the minimization problem (4.33) has $D^2 + 1$ unknown variables. This problem can be solved by the linear programming method previously outlined for the case $Q = \infty$. We need only to replace the system of linear inequalities (4.13)-(4.15) with the inequalities (4.33). We could also use the policy improvement algorithm to solve (4.33). The policy improvement algorithm for this case is similar to that of the infinite capacity case. In the following we simply state the changes necessary to accomplish this.

Step 2: Replace (4.16) with the following system of equality:

$$(4.34)\quad \begin{cases} \nu(x,y,\delta) = h\lambda^{-1}(x+y) - \lambda^{-1}\phi_\pi + p\nu(x+1,y,\delta) + q\nu(x,y+1,\delta) \\ \qquad\qquad\qquad\qquad \text{for } (x,y,\delta) \in \tilde{S}_\pi \\ \nu(x,y,\delta) = r - \mu\phi_\pi + h\mu(x+y-\delta y\wedge Q-\tilde{\delta}x\wedge Q) + R_\delta(x-\delta x\wedge Q, y-\tilde{\delta}y\wedge Q) \\ \qquad\qquad\qquad\qquad \text{for } (x,y,\delta) \in S_\pi . \end{cases}$$

Step 4: In (4.17) compute new values of S_π and $\tilde{S}_\pi$ using (4.32).

Step 5: Compute t_0 and t_1 using

$$(4.35)\quad \begin{cases} t_0 = \mu^{-1}\{\quad r \quad + h\mu(x+y-\delta y\wedge Q-\tilde{\delta}x\wedge Q) \\ \qquad\qquad + R_\delta(x-\delta x\wedge Q, y-\tilde{\delta}y\wedge Q) - \nu(x,y,\delta)\} \\ t_1 = \lambda\{h\lambda^{-1}(x+y) + p\nu(x+1,y,\delta) + q\nu(x,y+1,\delta) - \nu(x,y,\delta)\} . \end{cases}$$

All other steps remain the same as in the case of $Q = \infty$.

We conclude with the remark that both the linear program and policy improvement algorithm for the case $Q < \infty$ will give approximate results. However, if we make D very large, then the resulting error will be

small. One can also solve (4.34) for a sequence of increasing values of D and terminate the algorithm when further increase in D does not change the optimal policy. Some insight about the error in this procedure may be obtained by checking whether the last three terms of $R_\delta(0,0)$ in (4.30) and (4.31) is small compared to the first term of $R_\delta(0,0)$.

References

[1] Barnett, A. (1973). On operating a shuttle service, Networks, 3, 305-314.

[2] Deb, R. K. (1977). Optimal dispatching of a finite capacity shuttle, Technical Report No. 47, Department of Operations Research, Stanford University, Stanford, California.

[3] Deb, R. K. and R. F. Serfozo (1973). Optimal control of batch service queues, Advances in Applied Probability, 5, 340-361.

[4] Deb, R. K. (1976). Optimal control of batch service queues with switching cost, Advances in Applied Probability, 8, 177-194.

[5] Ignall, E. and P. Kolesar (1972). Operating characteristics of a simple shuttle under local dispatching rules, Operations Research, 20, 1077-1088.

[6] Ignall, E. and P. Kolesar (1974). Optimal dispatching of an infinite-capacity shuttle: Control at a single terminal, Operations Research, 22, 1008-1025.

[7] Ignall, E., R. Jagannathan and P. Kolesar (1976). Operating characteristics of an infinite capacity shuttle: Control at both terminals, Technical Report.

[8] Ross, S. M. (1970). Applied Probability Models with Optimization Applications, Holden Day, San Francisco.

[9] Saaty, T. L. (1961). Elements of Queueing Theory with Applications. McGraw-Hill Company, Inc., New York.

[10] Serfozo, R. E. (1978). Optimal control of random walks, Technical Report, Department of Industrial Engineering and Operations Research, Syracuse University, Syracuse, New York.

[11] Takacs, L. (1962). Introduction to Theory of Queues, Oxford University Press, New York.

[12] Widder, D. (1941). The Laplace Transform, Princeton University Press, Princeton, New Jersey.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 78	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) COMPUTATION OF THE OPTIMAL AVERAGE COST POLICY FOR THE TWO TERMINAL SHUTTLE		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) RAJAT K. DEB		8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0418
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Operations Research Stanford University Stanford, California		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (NR-047-061)
11. CONTROLLING OFFICE NAME AND ADDRESS OPERATIONS RESEARCH PROGRAM Code 434 Office of Naval Research Arlington, Virginia 22217		12. REPORT DATE APRIL 30, 1978
		13. NUMBER OF PAGES 30
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE, DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

This research was supported in part by National Science Foundation Grant ENG 75-14847 Department of Operations Research, Stanford University and issued as Technical Report No. 48

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Finite Capacity Shuttle, Bulk Queue, Markov Decision Process, Dynamic Programming, Optimal Control, Monotone Policies.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

SEE REVERSE SIDE

20. Abstract

COMPUTATION OF THE OPTIMAL AVERAGE COST POLICIES

FOR THE TWO TERMINAL SHUTTLE

Rajat K. Deb*
Stanford University, Stanford, California

In this paper we consider the problem of determining the optimal average cost policy for operating a shuttle between two terminals. The passengers arrive at each of the terminals according to Poisson processes and are transported by a single carrier with capacity $Q (\leq \infty)$ operating between the terminals. Under a fairly general cost structure, we show that the optimal average cost policy is monotone. Bounds are derived for the optimal control function and computational procedures for determining the optimal policy for both the finite and infinite capacity cases are presented.

< or = infinity

78/48